

Research Article

Domain-Oriented Subject Aware Model for Multimedia Data Retrieval

Lingling Zi,^{1,2} Junping Du,¹ and Qian Wang¹

¹ Beijing Key Laboratory of Intelligent Telecommunication Software and Multimedia, School of Computer Science, Beijing University of Posts and Telecommunications, Beijing 100876, China

² School of Electronic and Information Engineering, Liaoning Technical University, Huludao 125105, China

Correspondence should be addressed to Junping Du; junpingdu@126.com

Received 26 March 2013; Revised 22 May 2013; Accepted 23 May 2013

Academic Editor: Hua Li

Copyright © 2013 Lingling Zi et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the increment of the scale of internet information as well as the cross-correlation interaction, how to achieve accurate retrieval of multimedia data is an urgent question in terms of efficiently utilizing information resources. However, existing information retrieval approaches provide only limited capabilities to search multimedia data. In order to improve the ability of information retrieval, we propose a domain-oriented subject aware model by introducing three innovative improvements. Firstly, we propose the text-image feature mapping method based on the transfer learning to extract image semantics. Then we put forward the annotation document method to accomplish simultaneous retrieval of multimedia data. Lastly, we present subject aware graph to quantify the semantics of query requirements, which can customize query threshold to retrieve multimedia data. Conducted experiments show that our model obtained encouraging performance results.

1. Introduction

With the development of modern information technology, the manifestation of travel information has gradually changed from single text data to multimedia data. However, due to the continuing growth of tourism multimedia data and the fact that users are unable to express query requirements accurately, much time is spent on scanning and skimming through the results returned [1, 2], which means that the key problem to be addressed in information search is the development of a search model to guarantee the capability of understanding query requirements completely. The existing tourism information retrieval models are mostly keyword-based and therefore provide limited capabilities to capture user implicit query need. In face of this situation, information retrieval, as well as its related theories and technologies have been proposed nowadays. Nevertheless, these approaches exhibit a common limitation, which is the inability to take quantitatively semantic relations into account. In this paper, the previous problem can be addressed through the domain-oriented subject aware model (DSAM). This model will achieve the following objectives: (1) to develop a pattern that

enables unified multimedia data (i.e., text data and image data) in the tourism domain, (2) to analyze and quantify user implication requirements, and (3) to generate accurate multimedia search results for users. Through this model, the multimedia query results can be obtained in a precise and comprehensive way.

The development of DSAM involves many technologies, such as ontology, semantic search, and query expansion. Ontology is proposed for analyzing domain knowledge and used in all kinds of domains, especially in information retrieval [3–8]. For example, Setchi et al. [9] develop an image retrieval tool through ontological concepts, Chu et al. [10] construct a concept map learning system for education, and Dong et al. [11] propose a semantic service search engine for digital ecosystem. Meanwhile, as a knowledge representation form, ontology has been applied in the system development to provide implication query results, such as peer knowledge management system [12] and query-based ontology knowledge acquisition system [13]. In this paper, we are inspired by the idea of domain ontology and apply the definitions of concept and instance in the ontology to establish a subject aware graph in the tourism domain.

The semantic search technology [14–17] is also used in DSAM to capture the conceptualizations associated with the user query requirements. This technology is very popular in information retrieval [18], and many semantic search approaches have been proposed. For example, Hollink et al. [19] propose a method to exploit semantic information in the form of linked data. Bollegala et al. [20] describe empirical method to estimate semantic similarity using page counts and texts.

To obtain accurate and stable multimedia retrieval performance, we explore query expansion technique [21–23], which can be classified as local analysis, global analysis, and semantic dictionary method. In local analysis method, the expansion words are identified by using the most relevant articles which are associated the initial query [24]. In global analysis method, all the associated words or phrases of the entire document collection are used for correlation analysis, and the words associated with the highest degree of query word or phrase are added to new inquiries [25]. Finally, regarding the semantic dictionary method [26], Alejandra Segura et al. [27] focus the expansion on the use of domain ontology. In view of the features of these approaches, DSAM proposed can not only avoid all the words of the relevant calculation in the global analysis method as well as user participation feedback in the local analysis, but also cut down the cost of maintaining dictionary in the semantic dictionary method.

In conclusion, the novel contributions of this paper are the following: (1) we use the text mining technology and lots of text information to assist the knowledge learning of the image data and present text-image feature mapping method to extract image semantic. The advantage of our method is using relevant text information to assistantly generate the semantics of images, so as to improve the accuracy of image semantic annotation. (2) We propose the method of annotating documents to achieve the task of multimedia data fusion, including annotating creation and ranking of documents. This method can give more prominence to the important searching results and also capture a comprehensive understanding corresponding to user's query in a shorter time. (3) We propose the definition of subject aware graph (SAG) to quantify the semantics of the user query keywords. Furthermore, SAG contains three layers, that is, subject layer, concept layer, and instance layer. Meanwhile, the appropriate concepts and instances are organized rationally. In addition, we present the definition of awareness and its computing formulae for tackling the problem of measuring implicated query intention. And Awareness computations can be achieved using a thorough analysis of query requirements. As far as we know, this method has not been attempted in an information search system. (4) We present the implication of our model, including the information collection module, the index module, the subject aware expansion module, and the sorting and displaying module. DSAM explores the use of query threshold to support more accurate tourism multimedia search results, thereby improving the performance of retrieval.

The rest of the paper is structured as follows. Section 2 provides the concept of subject aware graph. Section 3 illustrates the implementation of our model. Section 4

presents experimental work to demonstrate the effectiveness of our model. Section 5 concludes the paper.

2. Subject Aware Graph

In this section, first we propose the concept of subject aware graph which is the foundation of Awareness. Then we elaborate the definition and calculation about awareness in order to obtain user implication query semantics. Last we demonstrate the application of awareness, which can be used in the DSAM implementation.

A subject aware graph consists of three parts: the subject layer containing subject nodes, the concept layer containing concept nodes, and the instance layer containing instance nodes. Three types of nodes are defined as follows.

Definition 1 (subject node). A subject node SN is in a 4-tuple form, where sid is the identity of SN, h is the level of this subject, n_c is the concept number associated with SN, and n_s is the number of child nodes of SN. Subject nodes are divided into two types, one is connection node (i.e., n_s is not zero) and the other is the leaf node (i.e., n_s is zero).

Definition 2 (concept node). A concept node CN is in a triple form, where cid is the identity of CN, $sort$ is the kind of CN (i.e., according to the concept property, $sort$ is divided into three categories, *basic concept*, *association concept*, and *comment concept*, resp.), and n_i is the instance number associated with CN.

Definition 3 (instance node). An instance node IN represents an instance of a concept associated with the given subject, with serial number used to identify IN.

According to the different types of nodes, we define awareness to quantify the semantics of the user query keywords, shown as follows.

Definition 4 (awareness). Awareness is a range of decimal $(0, 1]$, indicating the expansion degree of nodes in the SAG. Awareness includes three types, namely, subject awareness (SA), concept awareness (CA), and instance awareness (IA), which correspond to three layers of the SAG, respectively.

Subject awareness reflects the degree of subject concerned by people, and for calculating of SA, the following factors are considered. The first factor is h introduced in advance. The greater the level of SN, the less the contents of SN, so the smaller the value of SA. The second factor is n_s , and it is clear that the greater n_s is, the more dispersed its subject attention is and the less attention it attracts. The third factor is n_c , and furthermore, the larger the concept node number contained by SN is, the bigger the value of SA is. The last factor is the ratio of this subject resources denoted by this SN to total resources (P_s for short), and a higher ratio indicates that the subject is more attached by the people.

Taking all these factors, let SA be a list of weighted matrixes, namely, $SA = \{(m_1, w_1), (m_2, w_2), (m_3, w_3), (m_4, w_4)\}$ where $\sum_{i=1}^4 w_i = 1$. In this context, we define

matrixes as follows: $m_1 = f_1(h)$, $m_2 = 1/(n_s + 1)$, $m_3 = (n_c + 1)/(M_c + 1)$, $m_4 = \kappa_1 * P_s$, where $f_1(\mu) = (11 - \mu)/10$, $\kappa_1 = 10$ is an amplification constant and M_c is the maximum number of concepts contained by this SN.

Therefore, the SA with respect to a SN can be calculated with the following formula:

$$SA = \sum_{j=1}^4 m_j w_j, \quad (1)$$

where j ranges over all the matrixes in the description of SA.

For the computation of CA, we mainly consider two factors. The first factor is the ranking of concept type (denoted by r) whose order is the basic concept in the first place, the association concept in second place, and the comment concept in third place. The second factor is the instance number contained by the concept (denoted by n_i). This is because the former reflects the impact of concept type (i.e., the smaller the ranking number of the concept, the greater the CA of its concept), and the latter reflects the importance of instances (i.e., the more the instance number of the concept, the greater the CA of its concept). Based on previous two factors, we establish the CA formula as follows:

$$CA = f_1(r) f_2(n_i), \quad (2)$$

where function f_1 is consistent with SA formula, $f_2(n_i) = (n_i + 1)/(M_i + 1)$, where M_i is the maximum number of instances with any concept contained by the same subject.

Now, we present the formula of instance awareness as follows:

$$IA = \alpha * CA + \beta * \frac{n_i - n_{\min}}{n_{\max} - n_{\min}}, \quad (3)$$

where α and β are adjustment coefficients and satisfy $\alpha + \beta = 1$, n_i is the number of multimedia data contained by an instance, and n_{\min} and n_{\max} are the minimal and maximal numbers of multimedia data contained by any instance of the same subject, respectively. From previous equation, it can be seen that IA comprises two parts. The first part indicates the inheritance relationship between concept and instance; in other words, the higher CA is, the higher IA is. The second part indicates the attention degree of the instance through the linear conversion of multimedia data.

Finally, we elaborate the application of awareness. The idea of the Awareness calculations is to express the ambiguity of the query keywords input by users in the form of decimal. Returned comparison results CR is in a binary form $CR = \langle id, expansion \rangle$, where *expansion* represents expansion query keywords as user implicated subjects and *id* is its corresponding sequence number. Assuming that user query threshold is α ($0 < \alpha$) and subject node corresponding to user input query keywords is SK, we have the following comparison rules whose establishment principle is the larger the value of α (i.e. $\alpha > 1$) is, the wider the range of the subject is extended and the closer α is to 1 (i.e. $0 < \alpha < 1$), the more important the implicate keywords returned are to the given query keywords. Specifically, we have the following three application rules.

Rule 1. If $\alpha > 1$, implicit query keywords are subject nodes whose parent node is the same with SK and whose SA satisfies the following formula:

$$|SA - SA_{SK}| * h_1 < (\alpha - 1), \quad (4)$$

where SA_{SK} is the SA of SK and h_1 is an amplification factor. To facilitate the calculation, we change formula (4) to the following formula:

$$SA_{SK} + \frac{1 - \alpha}{h_1} < SA < SA_{SK} + \frac{\alpha - 1}{h_1}. \quad (5)$$

Rule 2. If the type of SK is leaf node under the condition of $0 < \alpha < 1$, then implicit query keywords are instance nodes which are related to the SK and satisfy $IA > \alpha$.

Rule 3. If the type of SK is connection node under the condition of $0 < \alpha < 1$, then implicit query keywords are subject nodes whose parent node is SK and the SA of these subject nodes satisfy the following formula:

$$\frac{SA - SA_{\min}}{SA_{\max} - SA_{\min}} > \alpha, \quad (6)$$

where SA_{\min} and SA_{\max} are, respectively, the minimal and maximal values of SA of subject nodes contained by the parent node SK. Similarly, we change formula (6) to the following formula:

$$SA > SA_{\min} + \alpha (SA_{\max} - SA_{\min}). \quad (7)$$

3. The Implication of DSAM

This proposed DSAM is not only able to capture accurately the user query intention, due to the fact that implication requirement is qualified through awareness calculations, but also to provide multifaceted tourism multimedia search results. The model architecture is presented in Figure 1, and it consists of four components, namely, information collection module, index module, subject aware expansion module, and sorting and displaying module. Firstly, the user enters query keywords and a query threshold into the query interface. Then, the subject aware expansion module generates an extended keyword set, and these keywords contained are delivered to the index module. Note that the index module achieves the function of creating indexes for annotation documents which have been established in the information collection module. Finally, the sorting and displaying module ranks the results returned from the index module and shows them through query interface.

3.1. Information Collection Module. Information collection module extracts semantics of multimedia resources, and the contents extracted are written in the label documents accordingly. Since different media types have different forms of resources, we unify them using the method of label documents at the semantic level. This module is specifically described as follows.

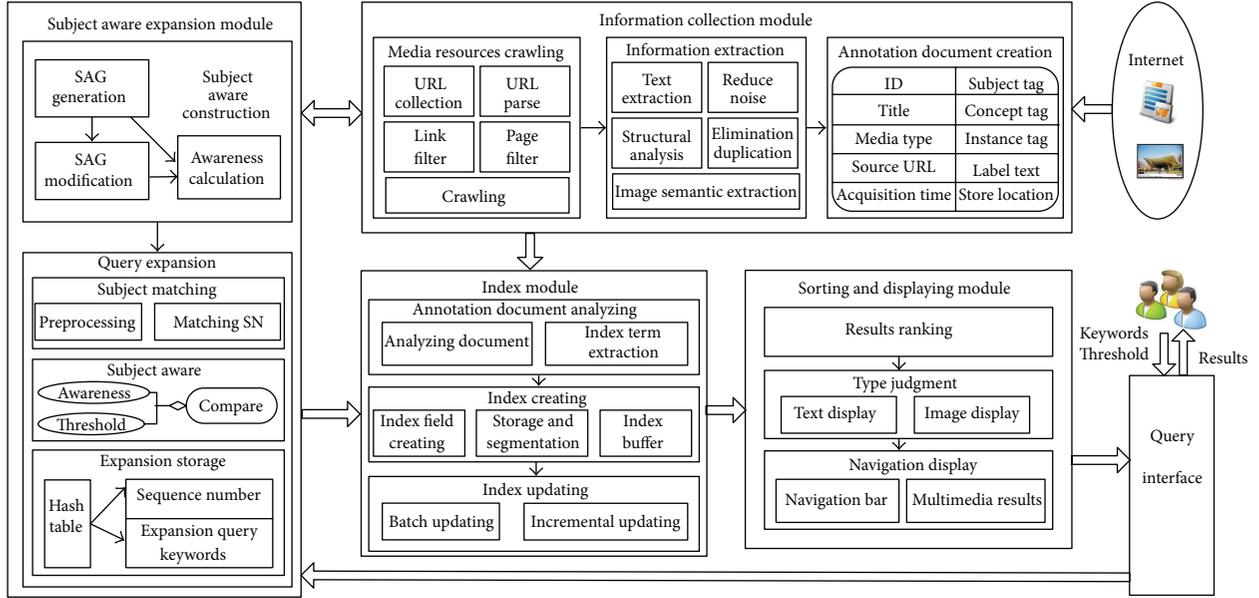


FIGURE 1: The architecture of DSAM.

(1) Media resources crawling: we use directional information collection method [28] to get the URL about tourism domain, and simultaneously, new URL can be produced by them. Then URL parsing is executed to detect the duplicate contents, and based on semantic analysis, the subject degree can be calculated. For the extracted links, we use the algorithm of extended metadata based on semantic analysis to calculate the subject correlated degree (see formula (8)), so as to implement link filters:

$$\text{sim}(u, v) = \frac{\sum_{k \in u \cap v} T_{ku} T_{kv}}{\sqrt{\sum_{k \in u} T_{kv}^2 \sum_{k \in v} T_{kv}^2}}, \quad (8)$$

where u represents the subject eigenvector, v represents the eigenvector of link texts, and T_{ku} is one of eigenvector terms in the feature vector space. On this basis, the subject evaluation value of collected pages can be conducted using keyword-based vector space model, shown as follows:

$$\text{NGD}(q_1, q_2) = \frac{\max\{\log y(q_1), \log y(q_2)\} - \log y(q_1, q_2)}{\log N - \min\{\log y(q_1), \log y(q_2)\}}, \quad (9)$$

where $y(q_i)$ represents the number of pages containing word q_i , N represents the total number of collected pages, and $y(q_i, q_j)$ represents the number of pages containing both word q_i and q_j . By excluding pages with low subject evaluation values, the accuracy of the collected subject pages can be improved. Finally, according to the results of the page filtering, web crawler automatically captures multimedia recourses (texts and images) and saves them in the corresponding database. In the process of crawling, the source URL and the acquisition time from every resource file are also recorded.

(2) Information extraction: firstly, the features of each resource file captured by the crawler are extracted as a vector

set. Then these features are converted into semantic information through the technique of structural analysis, noise reduction, duplicate content elimination, and text extraction [29–31]. Lastly, the semantic information is broken down into the subject tag, the concept tag, the instance tag, and label texts. Image semantic acquisition is a difficult point in multimedia information retrieval.

In order to accomplish the task of multimedia fusion, we use text-image feature mapping method based on the transfer learning [32, 33] to extract image semantic. The text data of each subject are modeled by using the latent Dirichlet allocation, and the corresponding discriminating text feature [34] can be captured by adopting the computation of information gain. The image data of each subject are modeled by utilizing the bag-of-visual-word mode [35, 36]. According to the feature distributions of the text data and the text-image cooccurrence data within the same subject, the feature distributions of the target images can be computed and then image semantic can be obtained, shown as follows:

$$P(g | s) = N_s \sum_{v \in V(s)} P(g | v, s, O) P(v | s, D), \quad (10)$$

where $P(g | s)$ denotes feature distributions of the target image within the subject s , $V(s)$ denotes the set of the most discriminating text feature contained by text set D , N_s denotes the normalization factor, $P(g | v, s, O)$ denotes the conditional probability distribution of the image feature, $P(v | s, D)$ denotes text feature distribution, and O denotes the set of text-image cooccurrence data.

(3) Annotation documents creation: we create annotation documents using the static mode, which is independent of the process of query. Its content is divided into three parts. The first part is document property information including the id and the title. The second part is resource collection information obtained from the step of media recourse crawling.

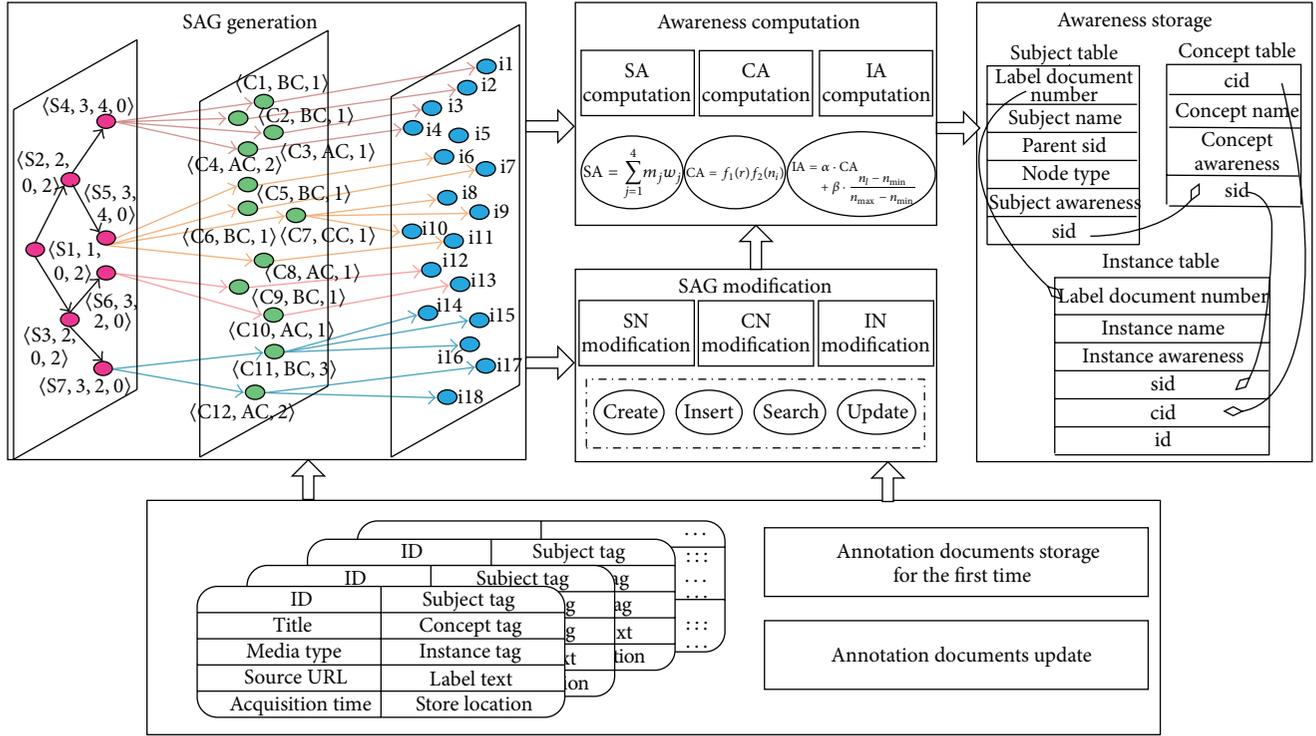


FIGURE 2: The process of subject aware construction.

The last part is document annotation information obtained from the step of information extraction. The creation of annotation documents lays the foundation for awareness computation which plays a role in quantifying user query requests.

3.2. Index Module. Aiming to quickly search information, we need to build up the index in the model. Index module can traverse all the annotation documents, extract index items, create index fields, and save them in the database. Specifically, the function of this module contains three parts. The first part is to analyze the contents of annotation documents obtained from the information collection module and extract index terms containing the title, the media type, the source URL, and label texts, which are used for establishing the corresponding index fields. On this basis, the second part is to create the inverted index whose form is denoted as $\langle k, \langle a_1, f_1, \langle p_{11}, p_{12}, \dots, p_{1f_1} \rangle \rangle, \dots, \langle a_i, f_i, \langle p_{i1}, p_{i2}, \dots, p_{if_i} \rangle \rangle, \dots, \langle a_k, f_k, \langle p_{k1}, p_{k2}, \dots, p_{kf_k} \rangle \rangle \rangle$, where k represents the number of the query words appearing in the annotation documents and a_i is the ID of the annotation document. Given the annotation document a_i , f_i is the term frequency of query word and $\langle p_{i1}, p_{i2}, \dots, p_{if_i} \rangle$ is its position list. Meanwhile, in the process of creating the index, we explore the techniques of storage and segmentation to obtain proper sets in different index fields. Also the cache technology can be used to improve the speed of index file creation. Since annotation documents need constant renewal and index files also need it correspondingly, the third part is to update in the manners of batch updating and incremental updating.

3.3. Subject Aware Expansion Module. The subject aware expansion module is the key component of the DSAM, including subject aware construction and query expansion. The former is the foundation of the latter.

3.3.1. Subject Aware Construction. The process of subject aware construction is shown in Figure 2. Firstly, we establish the SAG according to the contents of annotation documents and an overview of the process that follows in (Steps 1–4).

Step 1. Subject tags, concept tags, and instance tags are extracted from annotation document collection obtained from the information collection module.

Step 2. These tags are corresponding to the appropriate layers of SAG and new SN, CN, and IN can be simultaneously established. Particularly, the creation of SN includes traverse of the subject tree, search of parent nodes, insertion of the node, and record of the node information as well as the number increase of the annotation documents about this subject. Similarly, the creation of CN includes search of its SN, insertion of the node under this SN, and record of the node information (i.e., $c_{id}, sort, n_i$).

Step 3. According to SAG, the awareness (i.e., subject awareness, concept awareness, and instance awareness) can be computed (the awareness formula is described in Section 2).

Step 4. The computation results and related node information are stored in the subject table, concepts table, and instance table.

If new annotation document collections are obtained from the information collection module, SAG does not need to be created again, but the corresponding modifications include three cases shown as follows.

Case 1. SN modification: if the SN corresponds to the subject tag which is obtained in the new annotation document and has existed in the subject layer, then this SN can be found and its annotation document number increases. If not, a new SN needs to be created in the subject layer.

Case 2. CN modification: if the CN corresponds to the concept tag which is obtained in the new annotation document and has existed in the concept layer, there is nothing to do. If not, the SN related to this concept tag needs to be found and a new CN is inserted. Note that parameter n_c of this SN should be updated.

Case 3. IN modification: if the IN corresponds to the instance tag which is obtained in the new annotation document and has existed in the instance layer, then this IN can be found and its annotation document number increases. If not, the SN and the CN related to the instance tag need to be found and a new IN is inserted. Note that parameters of the SN and CN should be updated.

After the previous operations are completed, we recalculate the awareness and update the tables accordingly. Although the awareness computation needs to spend some time, it executes the task as a background process before

searching information and does not occupy the user's search time. Thereby it does not affect the efficiency of the system.

3.3.2. Query Expansion. When the user enters query keywords and the query threshold, a list of expansion keywords based on the calculations of subject aware expansion module can be obtained, and these expansion keywords reflect the potential user query intentions to some extent. Firstly, we carry out preprocessing (including null detection and Chinese word segmentation) according to the user query keywords. Then a SN can be matched in the SAG using the technique of word matching, and the application rules of awareness (see Section 2) can be performed. Lastly, the appropriate expansion lists returned are saved in the Hash table (for detailed algorithm, see Algorithm 1).

3.4. Sorting and Displaying Module. The sorting and displaying module consists of three parts: results ranking, media type judgment, and navigation display. We use the annotation sorting method to organize the searching results according to the correlation of the query expansion set and the annotation information. The specific processes are shown as follows.

Step 1. Calculate the correlation between expansion words and result records. Let $E = \{e_1, e_2, \dots, e_n\}$ be the extended word set. The degree of correlation between expansion word e_i and the annotation document, that is, $\text{Rank}(e_i, \text{label})$, is computed according to formula (11):

$$\text{Rank}(e_i, \text{label}) = \begin{cases} \sum_{j=1}^{\text{Occurrence}(e_i, \text{label})} \ln \frac{\text{Length}(\text{label})}{\text{Location}(e_i, j, \text{label})}, & \text{Occurrence}(e_i, \text{label}) > 0, \\ 0, & \text{Occurrence}(e_i, \text{label}) = 0, \end{cases} \quad (11)$$

where $\text{Length}(\text{label})$ represents the length of the annotation document; $\text{Occurrence}(e_i, \text{label})$ represents the frequency of e_i that occurs in the annotation document; $\text{Location}(e_i, j, \text{label})$ represents the location that e_i occurs in the annotation document. Then the correlation between extended word set and the annotation document, that is, label rank(E, label), is computed using the following formula:

$$\text{label rank}(E, \text{label}) = \sum_{i=1}^n \text{Rank}(e_i, \text{label}). \quad (12)$$

Step 2. Determine expansion degree of e_i , that is, ζ according to the position of the inverted index.

Step 3. Calculate the final correlation between E and annotation documents by using the following formula:

$$R(E, \text{label}) = \text{label rank}(E, \text{label}) \times \zeta. \quad (13)$$

Due to different contents of different media, the media type received from the field of index file needs to be judged, so as to determine the type of results displayed. Finally, multifaceted tourism information search results integrated

with text and image can be shown for users in the navigation view.

4. Experimental Results and Discussion

We have constructed subject aware system for users who query in Chinese inborn language. For the development of this system, we used Myeclipse 8.5 platform, MySQL 5.1, and a PC with Intel Core(TM) 2 Duo T6570 processor, 2.1 GHz and 4 GB of main memory. In this section, we collected 5000 multimedia objects as our experimental data set. These multimedia objects were from tourism sites on the Internet (such as Beijing travel, Sina web, Phoenix tourism and so on). The following parameters were used: $w_1 = 0.25$, $w_2 = 0.25$, $w_3 = 0.25$, $w_4 = 0.25$, $\alpha = 0.5$, and $\beta = 0.5$. Here we performed a comprehensive set of experiments to evaluate the performance of DSAM.

4.1. Evaluation of DSAM. In this experiment, we selected different numbers of multimedia objects to respond to eight query cases and then DSAM obtained the potential keywords (see Table 1). On this basis, we evaluated DSAM performance

Input: A subject aware graph G , user query threshold α , $\alpha > 0$, input query keywords QK
Output: Expansion result set CR

- (1) Initialize the result set CR to *null*.
- (2) Match between QK and the SN in the G to get SK ;
- (3) If ($\alpha > 1$) then search the corresponding results.
 - (a) Search all the SN whose parent node is the same with the parent node of SK and save them;
 - (b) Find the SN which satisfies the Rule 1, and rank the SN according to the difference between its SA and the SA of SK ;
 - (c) Save the sequence number of ranking as $CR.id$, the name of SN as $CR.expansion$;
- (4) search the corresponding results according to $\alpha \leq 1$
 - (a) If ($SK.n_s == null$) then
 - (i) find the IN in the G which satisfies the Rule 2, and rank the IN according to the IA of IN;
 - (ii) Save the sequence number of ranking as $CR.id$, the name of IN as $CR.expansion$;
 - (b) else
 - (i) Search all the SN in the G whose parent node is SK and save them to a set;
 - (ii) Find the SN which satisfies the Rule 3 from the set, and rank the SN according to the SA of SN;
 - (iii) Save the sequence number of ranking as $CR.id$, the name of SN as $CR.expansion$;
- (5) Return CR .

ALGORITHM 1: Subject aware query expansion algorithm.

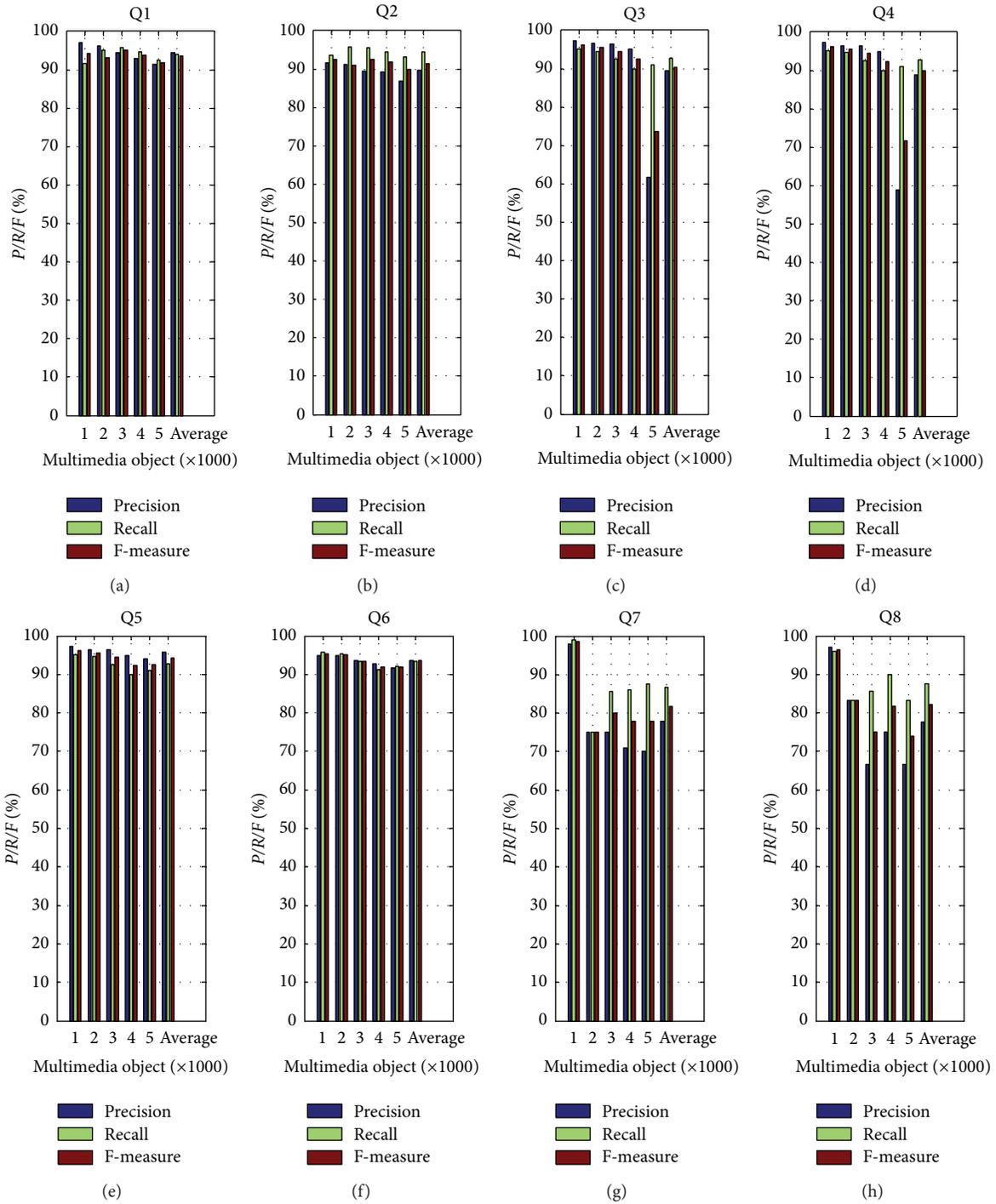
TABLE 1: The query case.

Query ID	Query keywords	Query threshold	Potential keywords
Q1	the Imperial palace	0.9	Hall of Supreme Harmony, Palace of Heavenly Purity, Palace of Earthly Tranquility,
Q2	the Imperial palace	0.2	Hall of Supreme Harmony, Palace of Heavenly Purity, Palace of Earthly Tranquility, Quanjude Restaurant, Hall of Harmony, Hall of Preserving Harmony, Beijing hotel, Prime Hotel, Wangfujing Grand hotel, Beijing International hotel, Dong-Lai-Shun restaurant, Tiananmen Square Fang Shan restaurant, Teahouse, Temple Fair.
Q3	Great wall badaling	1.5	Fragrant Hill
Q4	Great wall badaling	1.8	Fragrant Hill, Xiayunling
Q5	Natural scenery	0.7	Great Wall Badaling, Fragrant Hill, Xiayunling
Q6	Natural scenery	0.2	Great Wall Badaling, Fragrant Hill, Xiayunling, Fangshan Shidu, ChangpingHuyun, Mentougou, Miyun Longtan, Jingdong Grand Canyon, Yanqing Kangxi Grassland, Forest Canyon, Stream Waterfall, Grassland, Mountain
Q7	Old town	1.2	Temple
Q8	Old town	1.6	Temple, street scene

by Precision, Recall, and F -measure. Figure 3 shows $P/R/F$ results, respectively, under each query case. The average $P/R/F$ values corresponding to different numbers of multimedia objects are shown in Table 2. The results demonstrate that the performance of DSAM is relatively stable.

In order to further validate our model, we compare precision and recall values with Lucene. Figure 4 shows the comparison results in the case of the same query keywords under different numbers of multimedia objects. The following

two points can be seen: (1) with regard to the precision values, our results are slightly higher than those using Lucene in most cases. But when the number is 5000, the latter is higher than the former. This may be due to inaccuracy of image semantic. (2) With regard to recall values, our results are always obviously higher than those using Lucene. This is because that our model uses the subject aware query expansion algorithm to obtain more accurate query keywords. In conclusion, DSAM model has a relatively good performance.

FIGURE 3: $P/R/F$ results of different query case.TABLE 2: The average $P/R/F$ values.

The number of multimedia objects	The number of texts	The number of images	The average precision	The average recall	The average F -measure
1000	485	515	96.3	95.2	95.7
2000	960	1040	91.3	91.0	90.5
3000	1436	1564	88.6	91.7	89.9
4000	1909	2091	88.3	90.8	89.3
5000	2385	2615	77.7	90.2	82.9

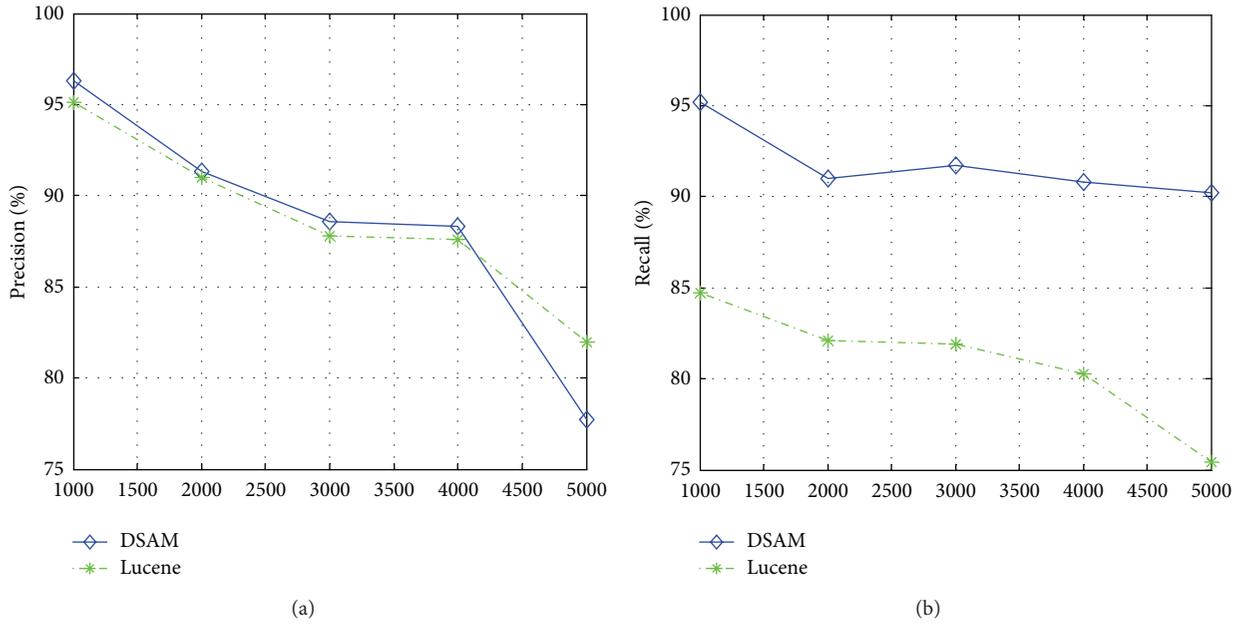


FIGURE 4: Comparison results between DSAM and Lucene. (a) shows comparison results of precision values under the different number of multimedia objects. (b) shows comparison results of recall values under the different number of multimedia objects.

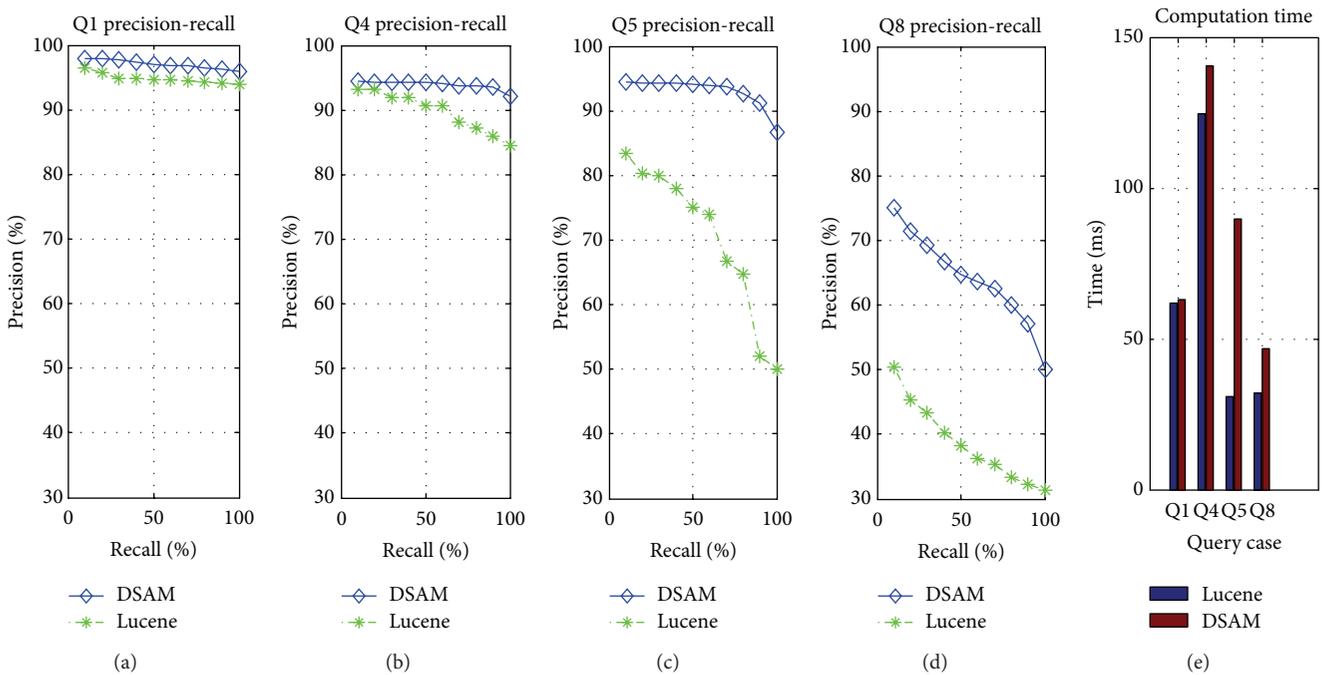
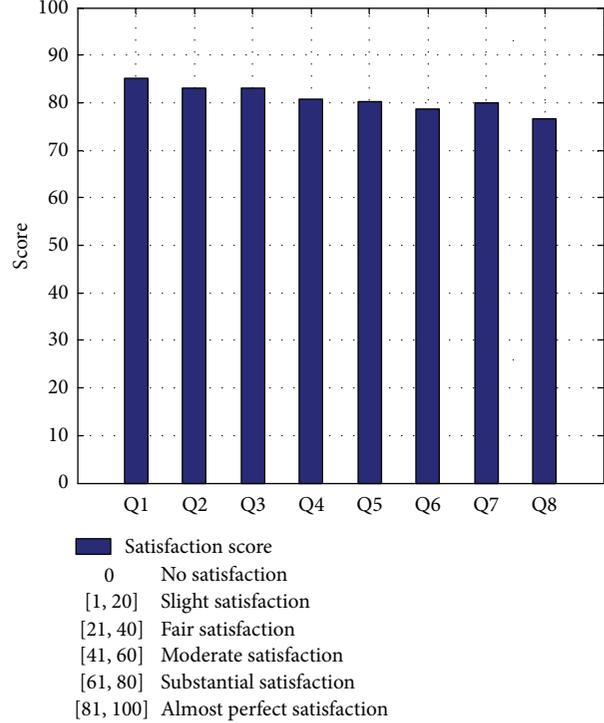
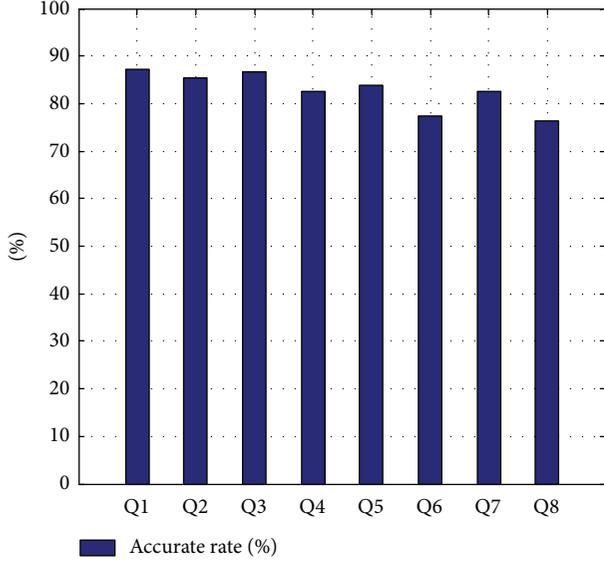


FIGURE 5: P-R curve between DSAM and Lucene and computation time.

Figures 5(a)–5(d) depict precision-recall curve for the four query cases (including Q1, Q4, Q5, and Q8), and Figure 5(e) records the time spent on achieving the previous query cases. Here we can see the following three points: (1) the precision-recall curve of DSAM is always above that of Lucene which means that our model is better than Lucene in terms of result coverage and result sort; (2) our model spends more time than that using Lucene in most cases (such as Q1, Q4, Q8) which is because that we need to retrieve more related query keywords. But the discrepancy is not very big;

(3) only for the query case Q5, due to the query keywords corresponding to connection node type, DSAM will produce comparatively more expending keywords and lead to time increase. In a word, our model uses less time to retrieve multimedia data in tourism domain.

We investigated the system performance evaluation from the perspective of the user with correct results provided by humans. For this reason, ten students from our department were asked to use this system. The volunteers entered the specified query keywords and thresholds (see Table 1) and



(a)

(b)

FIGURE 6: Performance evaluation by users.

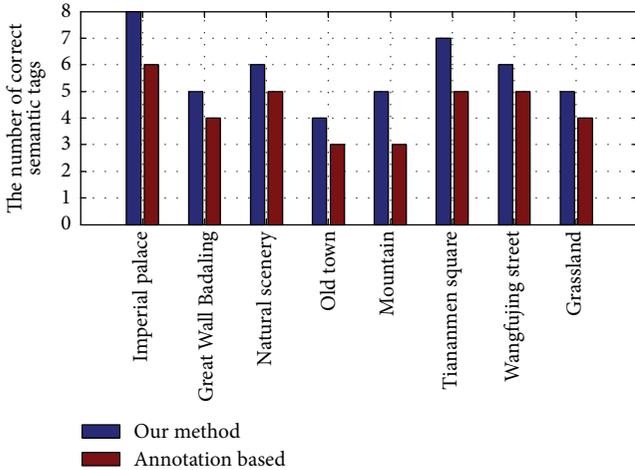


FIGURE 7: Comparison of the correctness of image semantic tags.

recorded ranking accuracy and satisfaction score according to the results returned. Figure 6(a) depicts the average ranking accurate rate of our survey, and we can see that the max accurate rate is 87.2%, the min 76.4%, and the average 82.8%. Note that from Q1 to Q4, the average accurate rate is 85.5%, while from Q5 to Q8, it is 80%. The average accurate rate of the former group is higher than the latter group which is because in the former group their query keywords corresponding to leaf node can lead to relatively clear query subject, while in the latter group, their query keywords corresponding to connection node can lead to

relatively broad query subject. Figure 6(b) summarizes the volunteers' average satisfaction score with regard to query results, where satisfaction standards of grading are shown on the right. The average satisfaction score is 80.9 which demonstrates that users are relatively satisfied with the query results. However, in our survey, there are also some cases of relatively low satisfaction score which is possibly because some multimedia objects are not marked accurately.

4.2. *Performance Comparison.* In order to evaluate the performance of the proposed text-image feature mapping method, we compare the accuracy of image semantic annotation with the annotation-based image retrieval method [22]. Figure 7 shows the obtained number of correct image semantic tags according to the different eight image themes in the field of tourism using the previous two methods. From this figure, we can observe that the proposed method obtains more correct semantic tags than the annotation-based image retrieval method. That is because compared method uses documents accompanying images to acquire image semantics. While our method utilizes the transfer learning technique to mine the feature mapping relationship between text information and image information, so as to obtain more correct image semantic tags.

Topic coverage and topic novelty are defined to evaluate the proposed annotation document method as shown in formula (14). The former reflects the comprehensiveness of query results and the latter embodies the ability to extend users' implicitly query intention. We compare topic coverage and topic novelty with Mediapedia [17]. Figure 8 shows the comparison results using the previous methods,

$$\text{Topic coverage} = \frac{\text{number of correctly returned topics}}{\text{number of relevant topics}}, \quad (14)$$

$$\text{Topic novelty} = \frac{\text{number of returned unknown topics}}{\text{number of known topics} + \text{number of returned unknown topics}}.$$

Now we will discuss Figure 8 from two aspects. On the one hand, the values of topic coverage using our method are generally higher than those of the comparison method. Since the average value of our method is 0.53 and the average value of Mediapedia is 0.47. This indicates that our method can obtain more correct query results. But as for the query cases Q1 and Q5, topic coverage values of our method are lower than those of Mediapedia. That is because the setting of our query threshold restricts the multimedia searching results. That is to say, only the very important information can be displayed. So the previous results demonstrate that our method can obtain more comprehensive multimedia information corresponding to the query requirement. On the other hand, the average value of topic novelty using our method is 0.3 and the average value of Mediapedia is 0.28, which indicates that our method can obtain more implicit information. For most of the query cases, topic novelty values of our method are higher than those of Mediapedia. That is because using our method, the related topic contents are retrieved and some of them are unknown to users. Only for the query cases Q1 and Q5, topic novelty values of our method are lower than those of Mediapedia, the reason of which is that the returned important contents are well known to users. Based on the previous two indicators, it can be seen that our method obtains a good effect on quantifying the semantics of user query.

P@10 evaluates the accuracy of the first ten returned results. Figure 9 shows the comparison results using Lucene, Semantic [4], and DSAM. From this figure, it can be clearly demonstrated that the proposed method outperforms the other two methods. Moreover, the following three points can be found. (1) The average value of P@10 obtained using our method is 0.575, while those of the other comparison methods are 0.35 and 0.45, respectively. It indicates that the search results obtained using DSAM are closer to the user's query requirement. (2) Since the core of Lucene is to use keyword search and that of Semantic is to expand context semantics through the ontology technique, some potential semantic information in the field of tourism cannot be found. While our method adopts the proposed subject aware graph to acquire the related subjects, concepts and instances correspond to the query, which is helpful for expanding user needs. So DSAM generally has a relatively high P@10 than the other comparison methods. (3) For the query case with the broad sense, such as Q6, the correlation of the query expansion set and the annotation information in DSAM is not very high. So the value of P@10 is lower than that of Semantic. In summary, this experiment proves that DSAM adopts the novel mode of customized query to acquire the desirable multimedia search results.

In the last experiment, we compare satisfaction scores of user evaluation for the obtained search results from different multimedia data set using three retrieval methods. And we invited another twenty students to do this experiment, as shown in Figure 10. The red curves represent the proposed method, the green curves Lucene method, and the blue curves Semantic method. Note that our curves are generally higher than the other retrieval methods. It indicates that more potential results including excellent images and texts are displayed using DSAM, which is recognized by more users. For each query case, satisfaction scores of user evaluation reduce accordingly with the increment of the number of multimedia data. That is because of that the precision rate and recall rate decrease is caused by multimedia data increment. In a word, the extensive experimental results show that the proposed DSAM outperforms the other comparison methods examined on subjective quality and quantitative measures.

5. Conclusions

This paper proposes a novel method of measuring user implicated query intention, and this method contains SAG establishment, awareness computation, and application. On this basis, we construct a subject aware multimedia retrieval model for tourism domain whose implementation has the following key points (1) in the information collection module, the text-image feature mapping and the annotation document methods are proposed to unify multimedia data; (2) in the index module, the inverted index according to the annotation document is established; (3) in the subject aware expansion module, a series of SAG and awareness operations are carried out. And the subject aware query expansion algorithm is presented to find the potential keywords; (5) in the sorting and displaying module, annotation sorting method is proposed, and multimedia query results are displayed in a precise and comprehensive way. To sum up, DSAM achieves accurate searching of tourism multimedia data through quantifying the relation between user query and the search results. Our experiments show that the proposed model can obtain encouraging performance in terms of objective evaluation and subjective evaluation. Future research will focus on improving the ranking accuracy of query results using ontology reasoning technology in order to provide better levels of tourism multimedia data.

Acknowledgments

This work was supported by the National Basic Research Program of China (973 Program) 2012CB821200

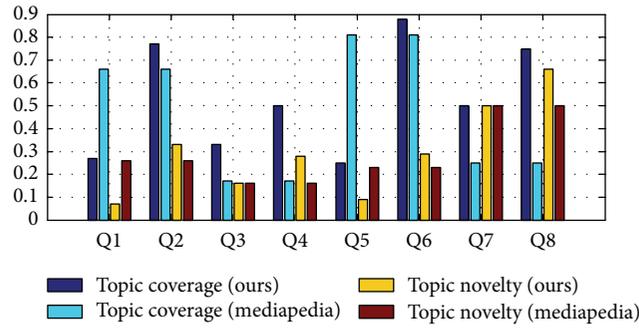


FIGURE 8: Comparison of the topic coverage and topic novelty.

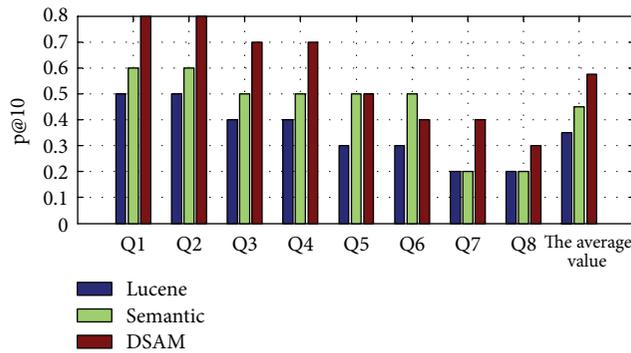


FIGURE 9: Comparison of P@10.

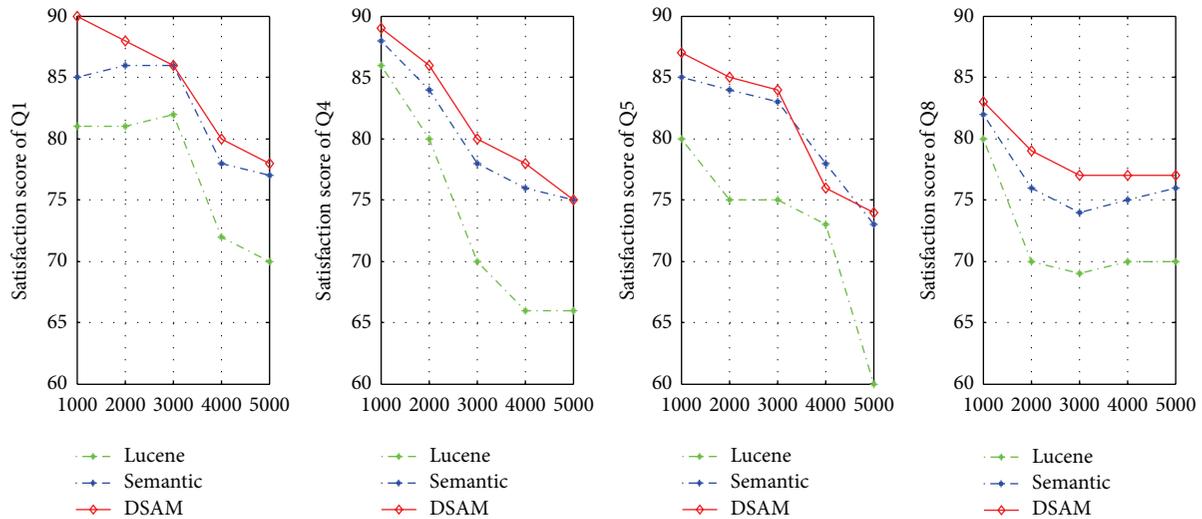


FIGURE 10: Comparison of satisfaction scores of user evaluation.

(2012CB821206), the National Natural Science Foundation of China (no. 91024001, no. 61070142), and the Beijing Natural Science Foundation (no. 4111002).

References

[1] Z. Xiang and B. Pan, "Travel queries on cities in the United States: implications for search engine marketing for tourist destinations," *Tourism Management*, vol. 32, no. 1, pp. 88–97, 2011.

[2] C. Carpineto and G. Romano, "A survey of automatic query expansion in information retrieval," *ACM Computing Surveys*, vol. 44, no. 1, article 1, 2012.

[3] A. Jimeno-Yepes, R. Berlanga-Llavori, and D. Rebbholz-Schuhmann, "Ontology refinement for improved information retrieval," *Information Processing and Management*, vol. 46, no. 4, pp. 426–435, 2010.

[4] M. Fernández, I. Cantador, V. López, D. Vallet, P. Castells, and E. Motta, "Semantically enhanced information retrieval: an ontology-based approach," *Journal of Web Semantics*, vol. 9, no.

- 4, pp. 434–452, 2011.
- [5] K. Robles, A. Fraga, J. Morato, and J. Llorens, “Towards an ontology-based retrieval of UML class diagrams,” *Information and Software Technology*, vol. 54, no. 1, pp. 72–86, 2012.
- [6] D. D. Kehagias, K. M. Giannoutakis, G. A. Gravvanis, and D. Tzovaras, “An ontology-based mechanism for automatic categorization of web services,” *Concurrency Computation Practice and Experience*, vol. 24, no. 3, pp. 214–236, 2012.
- [7] V.-S. Boonchom and N. Soonthornphisaj, “ATOB algorithm: an automatic ontology construction for Thai legal sentences retrieval,” *Journal of Information Science*, vol. 38, no. 1, pp. 37–51, 2012.
- [8] T. Kim and Y. S. Han, “Network behavior analysis simulation using ontology methodology,” *Information*, vol. 14, no. 10, pp. 3279–3287, 2011.
- [9] R. Setchi, Q. Tang, and I. Stankov, “Semantic-based information retrieval in support of concept design,” *Advanced Engineering Informatics*, vol. 25, no. 2, pp. 131–146, 2011.
- [10] K.-K. Chu, C.-I. Lee, and R.-S. Tsai, “Ontology technology to assist learners’ navigation in the concept map learning system,” *Expert Systems with Applications*, vol. 38, no. 9, pp. 11293–11299, 2011.
- [11] H. Dong, F. K. Hussain, and E. Chang, “A service search engine for the industrial digital ecosystems,” *IEEE Transactions on Industrial Electronics*, vol. 58, no. 6, pp. 2183–2196, 2011.
- [12] B. Qin, S. Wang, X. Du, Q. Chen, and Q. Wang, “Graph-based query rewriting for knowledge sharing between peer ontologies,” *Information Sciences*, vol. 178, no. 18, pp. 3525–3542, 2008.
- [13] T. I. Wang, T. C. Hsieh, K. H. Tsai, T. K. Chiu, and M. C. Lee, “Partially constructed knowledge for semantic query,” *Expert Systems with Applications*, vol. 36, no. 6, pp. 10168–10179, 2009.
- [14] J. Zhai, C. Yuan, Y. Chen, and J. Li, “Knowledge modeling and semantic retrieval of product data based on fuzzy ontology and sparql,” *Advanced Science Letters*, vol. 4, no. 4-5, pp. 1855–1859, 2011.
- [15] L. Kallipolitis, V. Karpis, and I. Karali, “Semantic search in the World News domain using automatically extracted metadata files,” *Knowledge-Based Systems*, vol. 27, pp. 38–50, 2012.
- [16] B. Yang and M. Mareboyana, “Similarity search in sensor networks using semantic-based caching,” *Journal of Network and Computer Applications*, vol. 35, no. 2, pp. 577–583, 2012.
- [17] H. Richang, Z. Zhengjun, G. Yue, C. Tat-Seng, and W. Xindong, “Multimedia encyclopedia construction by mining web knowledge,” *Signal Processing*, vol. 93, no. 8, pp. 2361–2368, 2013.
- [18] T. H. Cao and V. M. Ngo, “Semantic search by latent ontological features,” *New Generation Computing*, vol. 30, no. 1, pp. 53–71, 2012.
- [19] V. Hollink, T. Tsikrika, and A. P. De Vries, “Semantic search log analysis: a method and a study on professional image search,” *Journal of the American Society for Information Science and Technology*, vol. 62, no. 4, pp. 691–713, 2011.
- [20] D. Bollegala, Y. Matsuo, and M. Ishizuka, “A web search engine-based approach to measure semantic similarity between words,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 23, no. 7, pp. 977–990, 2011.
- [21] M. M. Rahman, S. K. Antani, and G. R. Thoma, “A query expansion framework in image retrieval domain based on local and global analysis,” *Information Processing and Management*, vol. 47, no. 5, pp. 676–691, 2011.
- [22] D. Kılınc and A. Alpkocak, “An expansion and reranking approach for annotation-based image retrieval from Web,” *Expert Systems with Applications*, vol. 38, no. 10, pp. 13121–13127, 2011.
- [23] A. L. Kaczmarek, “Interactive query expansion with the use of clustering-by-directions algorithm,” *IEEE Transactions on Industrial Electronics*, vol. 58, no. 8, pp. 3168–3173, 2011.
- [24] J.-H. Su, W.-J. Huang, P. S. Yu, and V. S. Tseng, “Efficient relevance feedback for content-based image retrieval by mining user navigation patterns,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 23, no. 3, pp. 360–372, 2011.
- [25] V. Jalali and M. R. M. Borujerdi, “Information retrieval with concept-based pseudo-relevance feedback in MEDLINE,” *Knowledge and Information Systems*, vol. 29, no. 1, pp. 237–248, 2011.
- [26] A. Ferrández, “Lexical and syntactic knowledge for information retrieval,” *Information Processing and Management*, vol. 47, no. 5, pp. 692–705, 2011.
- [27] N. Alejandra Segura, S.-S. Salvador-Sánchez, E. García-Barriocanal, and M. Prieto, “An empirical analysis of ontology-based query expansion for learning resource searches using MERLOT and the Gene ontology,” *Knowledge-Based Systems*, vol. 24, no. 1, pp. 119–133, 2011.
- [28] I. O. Ryzhov and W. B. Powell, “Information collection on a graph,” *Operations Research*, vol. 59, no. 1, pp. 188–201, 2011.
- [29] M.-H. Tsai, Y.-K. Chan, J.-S. Wang, S.-W. Guo, and J.-L. Wu, “Color-texture-based image retrieval system using Gaussian Markov random field model,” *Mathematical Problems in Engineering*, vol. 2009, Article ID 410243, 17 pages, 2009.
- [30] F. Su, G. Fang, and N. M. Kwok, “Adaptive colour feature identification in image for object tracking,” *Mathematical Problems in Engineering*, vol. 2012, Article ID 509597, 18 pages, 2012.
- [31] L. Kallipolitis, V. Karpis, and I. Karali, “Semantic search in the World News domain using automatically extracted metadata files,” *Knowledge-Based Systems*, vol. 27, pp. 38–50, 2012.
- [32] Z. Zhu, X. Zhu, Y. Ye, Y.-F. Guo, and X. Xue, “Transfer active learning,” in *Proceedings of the 20th ACM Conference on Information and Knowledge Management (CIKM ’11)*, pp. 2169–2172, October 2011.
- [33] S. J. Pan and Q. Yang, “A survey on transfer learning,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345–1359, 2010.
- [34] L. Wang, J. Zhang, and F. Zang, “An efficient feature extraction method, global between maximum and local within minimum, and its applications,” *Mathematical Problems in Engineering*, vol. 2011, Article ID 176058, 15 pages, 2011.
- [35] G. Qi, C. Aggarwal, and T. Huang, “Towards semantic knowledge propagation from text corpus to web images,” in *Proceedings of the 20th International Conference on World Wide Web*, pp. 297–306, 2011.
- [36] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, “Speeded-Up Robust Features (SURF),” *Computer Vision and Image Understanding*, vol. 110, no. 3, pp. 346–359, 2008.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

